

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 10-301720

(43)Date of publication of application : 13.11.1998

(51)Int.Cl.

G06F 3/06

G06F 3/06

G06F 3/06

G06F 12/16

G11B 19/02

(21)Application number : 09-107360

(71)Applicant : NEC IBARAKI LTD

(22)Date of filing : 24.04.1997

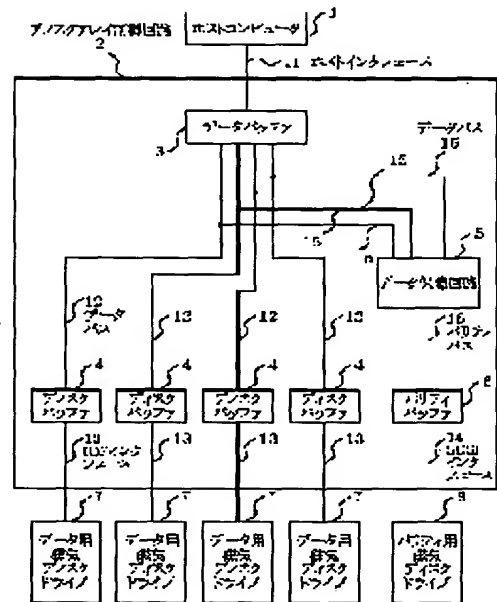
(72)Inventor : EDA YOSHINORI

(54) DISK ARRAY DEVICE

(57)Abstract:

PROBLEM TO BE SOLVED: To provide an inexpensive disk array device without lowering reliability by generating the CRC of respective magnetic disk drives in a data security circuit and storing it in the magnetic disk drive for a parity at the time of writing data to the magnetic disk drive for the data.

SOLUTION: The disk array of RAID is constituted of four magnetic disk drives 7 for the data and one magnetic disk drive 8 for the parity. The plural magnetic disk drives of the IDE interface 13 with low reliability since the CRC can not be added though inexpensive are used as the magnetic disk drives 7 for the data, only the magnetic disk drive 8 for the parity is turned to the magnetic disk drive of the SCSI interface 14 with high reliability though expensive and the disk array is constituted. Then, the CRC of the respective magnetic disk drives 7 for the data is stored in the magnetic disk drive 8 for the parity of the SCSI interface 14.



(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平10-301720

(43) 公開日 平成10年(1998)11月13日

(51) Int.Cl.⁸
G 0 6 F 3/06
12/16
G 1 1 B 19/02

識別記号
3 0 5
3 0 2
5 4 0
3 2 0
5 0 1

F I
G 0 6 F 3/06
12/16
G 1 1 B 19/02
3 0 5 C
3 0 2 A
5 4 0
3 2 0 L
5 0 1 F

審査請求 有 請求項の数 7 O L (全 5 頁)

(21) 出願番号 特願平9-107360

(22) 出願日 平成9年(1997)4月24日

(71) 出願人 000119793

茨城日本電気株式会社

茨城県真壁郡関城町関館字大茶367-2

(72) 発明者 江田 由則

茨城県真壁郡関城町関館字大茶367の2

茨城日本電気株式会社内

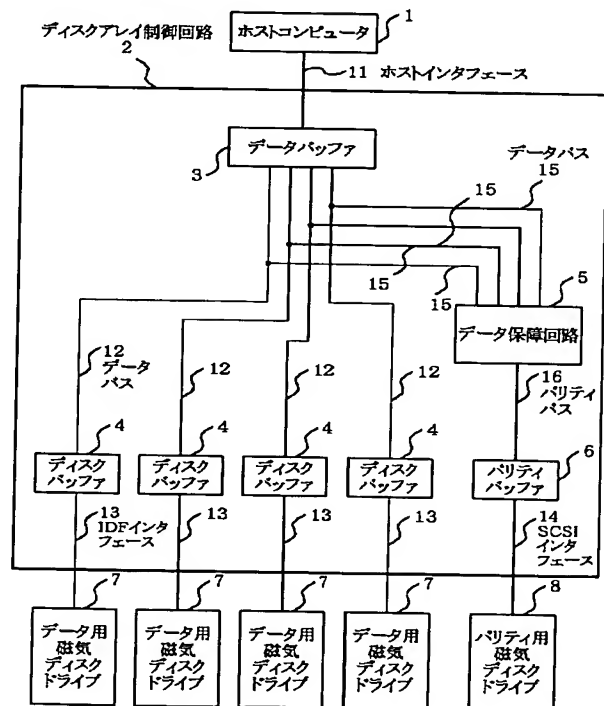
(74) 代理人 弁理士 京本 直樹 (外2名)

(54) 【発明の名称】 ディスクアレイ装置

(57) 【要約】

【課題】 信頼性を低下させずに安価なディスクアレイ装置を実現する。

【解決手段】 I D E インタフェースの複数台の磁気ディスクドライブをデータ用磁気ディスクドライブとして使用し、パリティ用磁気ディスクドライブのみを S C S I インタフェースの磁気ディスクドライブとしてディスクアレイを構成し、データ用磁気ディスクドライブにデータを書き込むとき、各磁気ディスクドライブの C R C をデータ保障回路で生成してパリティ用磁気ディスクドライブに格納しておく。



【特許請求の範囲】

【請求項1】 ホストコンピュータとの間でホストインタフェースを介して授受を行うデータを一時的に格納するデータバッファと、前記データバッファのデータを分割して格納する複数のディスクバッファと、前記複数のディスクバッファがそれぞれ対応しているデータ用磁気ディスクドライブとの間でデータの転送を行うときにパリティの生成およびパリティチェックならびに前記複数のデータ用磁気ディスクドライブに対応するチェックフレームの生成およびフレームチェックを行うデータ保障回路と、前記データ保障回路の生成したパリティおよびCRCを格納するパリティバッファとを有するディスクアレイ制御回路と、

I D Eインタフェースを有し前記複数のディスクバッファのそれぞれに接続された前記複数のデータ用磁気ディスクドライブと、

S C S Iインタフェースを有し前記パリティバッファに接続されたパリティ用磁気ディスクドライブとを備えることを特徴とするディスクアレイ装置。

【請求項2】 前記ディスクバッファおよび前記データ用磁気ディスクドライブの数を2個としたことを特徴とする請求項1記載のディスクアレイ装置。

【請求項3】 前記ディスクバッファおよび前記データ用磁気ディスクドライブの数を3個としたことを特徴とする請求項1記載のディスクアレイ装置。

【請求項4】 前記ディスクバッファおよび前記データ用磁気ディスクドライブの数を4個としたことを特徴とする請求項1記載のディスクアレイ装置。

【請求項5】 1セクタ当りのデータ長を512バイトとした前記データ用磁気ディスクドライブを備えることを特徴とする請求項1または請求項2または請求項3または請求項4記載のディスクアレイ装置。

【請求項6】 1セクタ当りのデータ長を520バイトとした前記パリティ用磁気ディスクドライブを備えることを特徴とする請求項1または請求項2または請求項3または請求項4または請求項5記載のディスクアレイ装置。

【請求項7】 1セクタ当りのデータ長を528バイトとした前記パリティ用磁気ディスクドライブを備えることを特徴とする請求項1または請求項2または請求項3または請求項4または請求項5記載のディスクアレイ装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】 本発明は、複数のデータ用磁気ディスクドライブと、それらを接続してホストコンピュータとの間でデータの授受を行うディスクアレイ制御回路とを備えるディスクアレイ装置に関する。

【0002】

【従来の技術】 従来のディスクアレイ装置は、ホストコ

ンピュータと接続するためのインタフェースとして、S C S Iインタフェース (Small Computer System Interface) を有しており、このため、ディスクアレイ制御回路とデータ用磁気ディスクドライブとを接続するためのインタフェースとしても、S C S Iインタフェースを採用している (例えば特開平8-137630号公報参照)。

【0003】

【発明が解決しようとする課題】 上述したように、従来のディスクアレイ装置は、ディスクアレイ制御回路とデータ用磁気ディスクドライブとを接続するためのインタフェースとしてS C S Iインタフェースを採用しているため、パーソナルコンピュータ用として広く普及している安価なI D E (Integrated Device (Drive) Electronics) インタフェースの磁気ディスクドライブを採用することができないという欠点を有している。

【0004】 I D Eインタフェースの磁気ディスクドライブは、1セクタ当りのデータ長を512バイトに固定されているため、各セクタ毎にC R C (Cyclic Redundancy Check) を付加することができず、従って信頼性が低いものである。これに対して、S C S Iインタフェースの磁気ディスクドライブは、1セクタ当りのデータ長を520バイトまたは528バイトとすることができるため、C R Cを付加して信頼性を向上させることができる。

【0005】 本発明の目的は、上述のような従来のディスクアレイ装置の欠点を解消して、安価なI D Eインタフェースの磁気ディスクドライブを接続し、しかもC R Cを付加できるように信頼性を向上させることができるディスクアレイ装置を提供することにある。

【0006】

【課題を解決するための手段】 本発明のディスクアレイ装置は、ホストコンピュータとの間でホストインタフェースを介して授受を行うデータを一時的に格納するデータバッファと、前記データバッファのデータを分割して格納する複数のディスクバッファと、前記複数のディスクバッファがそれぞれ対応しているデータ用磁気ディスクドライブとの間でデータの転送を行うときにパリティの生成およびパリティチェックならびに前記複数のデータ用磁気ディスクドライブに対応するチェックフレームの生成およびフレームチェックを行うデータ保障回路と、前記データ保障回路の生成したパリティおよびC R Cを格納するパリティバッファとを有するディスクアレイ制御回路と、I D Eインタフェースを有し前記複数のディスクバッファのそれぞれに接続された前記複数のデータ用磁気ディスクドライブと、S C S Iインタフェースを有し前記パリティバッファに接続されたパリティ用磁気ディスクドライブとを備えるものであり、特に、前記ディスクバッファおよび前記データ用磁気ディスクドライブの数を2個または3個または4個とし、

また、前記データ用磁気ディスクドライブの1セクタ当りのデータ長を512バイトとし、また、前記パリティ用磁気ディスクドライブの1セクタ当りのデータ長を520バイトまたは528バイトとしたとしたものである。

【0007】

【発明の実施の形態】次に、本発明の実施の形態について図面を参照して説明する。

【0008】図1は本発明の一実施形態を示すブロック図である。

【0009】図1において、ホストコンピュータ1とディスクアレイ制御回路2とは、ホストインタフェース11を介して接続されてデータの授受を行う。

【0010】ディスクアレイ制御回路2は、1個のデータバッファ3と、1個のデータ保障回路5と、4個のディスクバッファ4と、1個のパリティバッファ6とを有している。

【0011】データバッファ3は、ホストコンピュータ1とホストインタフェース11を介して接続されており、ホストコンピュータ1との間で授受を行うデータを一時的に格納する。データバッファ3には、データ保障回路5と4個のディスクバッファ4とが接続されている。

【0012】4個のディスクバッファ4は、それぞれデータバス12を介してデータバッファ3と接続されており、またそれぞれIDEインタフェース13を介してデータ用磁気ディスクドライブ7と接続されており、データバッファ3に格納されているデータを4分割したデータの1つのデータを格納する。

【0013】データ保障回路5は、データバス15を介してデータバッファ3およびディスクバッファ4と接続されており、またパリティバス16を介してパリティバッファ6とも接続されている。データ保障回路5は、データバッファ3のデータからパリティを生成してパリティバッファ6に出力し、4個のディスクバッファ4のそれぞれについてCRCを生成してパリティバッファ6に出力する。またディスクバッファ4およびパリティバッファ6からデータを入力してパリティチェックを行うと共に、CRCチェックも行う。

【0014】パリティバッファ6は、パリティバス16を介してデータ保障回路5と接続すると共に、SCSIインタフェース14を介してパリティ用磁気ディスクドライブ8と接続されている。

【0015】データ用磁気ディスクドライブ7は、1セクタ当りのデータ長を512バイトとしたIDEインタフェースの磁気ディスクドライブで、ディスクバッファ4とはIDEインタフェース13を介して接続されており、対応するディスクバッファ4から送られてきたデータを書き込んだり、ディスクバッファ4に格納するデータを読み出したりする。

【0016】パリティ用磁気ディスクドライブ8は、1セクタ当りのデータ長を528バイト（データが512バイト、CRCが16バイト）としたSCSIインタフェースの磁気ディスクドライブで、パリティバッファ6とSCSIインタフェース14を介して接続されている。

【0017】4台のデータ用磁気ディスクドライブ7と1台のパリティ用磁気ディスクドライブ8とで、RAID3のディスクアレイを構成している。

【0018】次に、上述のように構成したディスクアレイ装置の動作について説明する。

【0019】ホストコンピュータ1からディスクアレイ制御回路2に送られてきたデータは、ホストインタフェース11を介してデータバッファ3に入力し、そこに一時的に格納され、さらに4分割されてそれぞれデータバス12を介して4個のディスクバッファ4に格納される。データバッファ3に格納されたデータは、これと同時にデータ保障回路5にも送られる。データ保障回路5は、入力したデータについてパリティデータ（パリティ）を生成し、生成したパリティをパリティバッファ6に送る。パリティバッファ6は、入力したパリティを格納する。このときデータ保障回路5はまた、4台のデータ用磁気ディスクドライブ7のそれぞれについて、CRCを生成する。

【0020】次に、4個のディスクバッファ4から、それぞれに接続しているデータ用磁気ディスクドライブ7にデータを送ってそれを格納させると同時に、パリティバッファ6からパリティをパリティ用磁気ディスクドライブ8に送ってそれを格納させた後、4台のデータ用磁気ディスクドライブ7および1台のパリティ用磁気ディスクドライブ8のCRCを順次にパリティ用磁気ディスクドライブ8に送って格納させる。

【0021】次に、データの読み出しのときの動作について説明する。

【0022】4台のデータ用磁気ディスクドライブ7から読み出されたデータは、それぞれ対応するディスクバッファ4に格納される。このとき同時に、パリティ用磁気ディスクドライブ8からもパリティデータが読み出されてパリティバッファ6に格納される。

【0023】4個のディスクバッファ4に格納された読み出しデータは、それぞれ対応するデータバス12を経由してデータバッファ3に格納されると同時に、データ保障回路5にも入力する。このとき同時に、パリティバッファ6に格納されているパリティデータもデータ保障回路5にも入力する。

【0024】データ保障回路5は、入力した4個のディスクバッファ4からのデータおよびパリティデータから、4台のデータ用磁気ディスクドライブ7およびパリティ用磁気ディスクドライブ8の各ドライブ毎にCRCを生成する。

【0025】次に、データバッファ3に格納されているデータを、ホストインタフェース11を介してホストコンピュータ1に転送する。

【0026】4台のデータ用磁気ディスクドライブ7からのそれぞれ512バイトのデータをホストコンピュータ1に転送した後、パリティ用磁気ディスクドライブ8から読み出した5台の磁気ディスクドライブ（4台のデータ用磁気ディスクドライブ7と1台のパリティ用磁気ディスクドライブ8）に対応するCRCと、データ保障回路5で生成したCRCとを比較して4台のデータ用磁気ディスクドライブ7から読み出したデータの妥当性をチェックする。もしこのときエラーがあれば、ホストコンピュータ1にそのことを通知する。

【0027】次に、4台のデータ用磁気ディスクドライブ7の中の1台のデータ用磁気ディスクドライブ7が故障した場合のデータの読み出し動作について説明する。

【0028】故障した1台のデータ用磁気ディスクドライブ7を除く3台のデータ用磁気ディスクドライブ7から読み出されたデータは、それぞれ対応するディスクバッファ4に格納され、更にデータバッファ3に格納されると共に、データ保障回路5に入力する。このとき同時に、パリティ用磁気ディスクドライブ8からもパリティデータが読み出されてパリティバッファ6に格納され、更にデータ保障回路5に入力する。

【0029】データ保障回路5は、入力した3台のデータ用磁気ディスクドライブ7からデータとパリティ用磁気ディスクドライブ8からのパリティデータとにより、故障した1台のデータ用磁気ディスクドライブ7のデータを復元し、データバス15を介してデータバッファ3に格納する。データバッファ3は、すべてのデータが揃った時点でホストコンピュータ1に転送する。

【0030】このようにして、安価ではあるがCRCを付加できなために信頼性が劣るIDEインタフェースの複数台の磁気ディスクドライブをデータ用磁気ディスクドライブとして使用し、パリティ用磁気ディスクドライブのみを高価ではあるが信頼性が高いSCSIインタフェースの磁気ディスクドライブとしてディスクアレイを構成し、各データ用磁気ディスクドライブのCRCをS

CSIインタフェースのパリティ用磁気ディスクドライブに格納しておくことにより、信頼性を低下させずに安価なディスクアレイ装置を実現することができる。

【0031】上述の実施形態は、データ用磁気ディスクドライブの数を4台としたものであるが、データ用磁気ディスクドライブの数は、2台または3台または5台以上とすることができる。また、パリティ用磁気ディスクドライブの1セクタ当りのデータ長を、520バイト（データが512バイト、CRCが8バイト）とすることもできる。

【0032】

【発明の効果】以上説明したように、本発明のディスクアレイ装置は、IDEインタフェースの複数台の磁気ディスクドライブをデータ用磁気ディスクドライブとして使用し、パリティ用磁気ディスクドライブのみをSCSIインタフェースの磁気ディスクドライブとしてディスクアレイを構成し、データ用磁気ディスクドライブにデータを書き込むとき、各磁気ディスクドライブのCRCをデータ保障回路で生成してパリティ用磁気ディスクドライブに格納しておくことにより、信頼性を低下させずに安価なディスクアレイ装置を実現することが可能になるという効果がある。

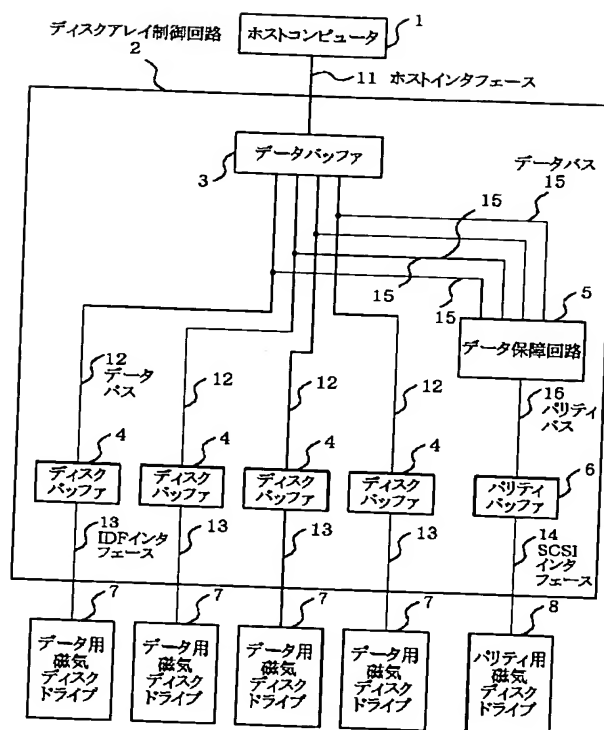
【図面の簡単な説明】

【図1】本発明の一実施形態を示すブロック図である。

【符号の説明】

- 1 ホストコンピュータ
- 2 ディスクアレイ制御回路
- 3 データバッファ
- 4 ディスクバッファ
- 5 データ保障回路
- 6 パリティバッファ
- 7 データ用磁気ディスクドライブ
- 8 パリティ用磁気ディスクドライブ
- 11 ホストインタフェース
- 12・15 データバス
- 13 IDEインタフェース
- 14 SCSIインタフェース
- 16 パリティバス

【圖 1】



United States Patent
Mattson

5,434,992
July 18, 1995

Method and means for dynamically partitioning cache into a global and data type subcache hierarchy from a real time reference trace

Abstract

A method and means is disclosed for dynamically partitioning an LRU cache partitioned into a global cache storing referenced objects of k different data types and k local caches storing objects of a single type. Referenced objects are stored in the MRU position of the global cache and overflow is managed by destaging the LRU object from the global to the local cache having the same data type. Dynamic partitioning is accomplished by recursively creating and maintaining from a trace of objects an LRU list of referenced objects and associated data structures for each subcache, creating and maintaining a multi-planar array of partition distribution data from the lists and the trace as a collection of all possible of maximum and minimum subcache sizing, optimally resizing the subcache partitions by applying a dynamic programming heuristic to the multiplanar array, and readjusting the partitions accordingly.

Inventors: Mattson; Richard L. (San Jose, CA)
Assignee: International Business Machines Corporation (Armonk, NY)
Appl. No.: 940560
Filed: September 4, 1992

Current U.S. Class:

711/119; 711/129; 711/136; 711/173

Intern'l Class:

G06F 012/00; G06F 013/00

Field of Search: